

IN THE CLAIMS:

Claims 1, 7-10, 16, 18, 20, 26, 28, 29, 31-33, 35, 36, and 39-42 are amended herein.

Claim 13 is canceled. All pending claims and their present status are produced below.

1. (Currently Amended) An apparatus for direct annotation of objects, the apparatus comprising:
 - a display device for displaying one or more images;
 - an audio input device for receiving an audio input; and
 - a direct annotation creation module coupled to receive an input audio signal from the audio input device and to receive a reference to a location within an image from the display device, the direct annotation creation module, in response to receiving the audio input audio signal and the reference to the location within the image, automatically creating an annotation object, independent from the image, that associates the input audio signal with the image location, and the direct annotation creation module automatically terminating a recording of the input audio signal based on a predetermined audio level.
2. (Original) The apparatus of claim 1 further comprising an annotation display module coupled to the direct annotation creation module, the annotation display module generating symbols or text representing the annotation objects.
3. (Original) The apparatus of claim 1 further comprising an annotation audio output module coupled to the direct annotation creation module, the annotation audio output module

generating audio output in response to user selection of an annotation symbol representing an annotation object.

4. (Original) The apparatus of claim 1 further comprising:
an audio vocabulary storage for storing a plurality of audio signals and corresponding text strings;
an audio vocabulary comparison module coupled to the audio input device, the audio vocabulary storage and the direct annotation creation module, the audio vocabulary comparison module receiving audio input and finding a corresponding text string that matches the audio input; and
wherein the direct annotation creation module uses text strings found by the audio vocabulary comparison module to create the audio annotation.

5. (Original) The apparatus of claim 1 further comprising:
an audio vocabulary storage for storing a plurality of audio signals and corresponding text strings;
a dynamic vocabulary updating module coupled to the audio vocabulary storage and the audio input device, the dynamic vocabulary updating module for displaying an interface to create a new entry in the audio vocabulary storage, the dynamic vocabulary updating module receiving an audio input and a text string and creating the new entry in the audio vocabulary storage.

6. (Original) The apparatus of claim 1 further comprising a media object cache for storing media and annotation objects.

7. (Currently Amended) An apparatus for direct annotation of objects for use with a system for storing, accessing, and presenting objects such as video objects, text objects, audio objects, or image objects, the apparatus comprising:

a direct annotation creation module coupled to receive an input audio signal and a reference to a location within an image, the direct annotation creation module, in response to receiving the input audio signal or the reference to the location within the image, automatically creating an annotation object, independent of the image, that associates a symbol or text with the image location, and the direct annotation creation module automatically terminating a recording of the input audio signal based on a predetermined audio level; and

an annotation display module coupled to the direct annotation creation module, the annotation display module generating the symbol or text representing the annotation object on a display device.

8. (Currently Amended) An apparatus for direct annotation of objects for use with a system for storing, accessing, and presenting objects such as video objects, text objects, audio objects, or image objects, the apparatus comprising:

a direct annotation creation module coupled to receive an input audio signal and a reference to a location within an image, the direct annotation creation module, in response to receiving the input audio signal or the reference to the location within the image, automatically creating an annotation object, independent of the image, that associates the input audio signal and the image location, and the direct annotation creation module automatically terminating a recording of

the input audio signal based on a predetermined audio level, the annotation object including at least an audio input field, an image reference field, and an annotation location field; and
an annotation audio output module coupled to the direct annotation creation module, the annotation audio output module generating audio output in response to user selection of an annotation symbol representing the annotation object.

9. (Currently Amended) An apparatus for direct annotation of objects, the apparatus comprising:

a media object storage for storing media and annotation objects; and
a direct annotation creation module coupled to receive an input audio signal and a reference to a location within an image, the direct annotation creation module, in response to receiving the input audio signal or the reference to the location within the image, automatically creating an annotation object, independent of the image, that associates the input audio signal and the image location, and the direct annotation creation module automatically terminating a recording of the input audio signal based on a predetermined audio level, and [,,] the direct annotation creation module storing the audio annotation in the media object storage.

10. (Currently Amended) A computer implemented method for direct annotation of objects, the method comprising the steps of:

displaying an image;

receiving audio input;

detecting selection of [[an]] a location within the image; and

creating an annotation object, independent of the selected image, between the selected image location and the audio input, the annotation object including at least an audio input field, an image reference field, and an annotation location field,
the creating step occurring automatically in response to the receiving or
detecting steps and including automatically terminating a recording of the
audio input based on a predetermined audio level.

11. (Original) The method of claim 10, wherein the step of displaying is performed before or simultaneously with the step of receiving.

12. (Original) The method of claim 10, wherein the step of receiving is performed before or simultaneously with the step of displaying.

13. (Canceled)

14. (Original) The method of claim 10, further comprising the step of displaying a visual notation that the image has an annotation.

15. (Original) The method of claim 14, wherein the visual notation is text or a symbol.

16. (Currently Amended) The method of claim 10, wherein the step of creating an annotation object includes ~~creating an annotation object~~ and storing the annotation object in an object storage.

17. (Original) The method of claim 10, further comprising the step of recording the audio input received.

18. (Currently Amended) The method of claim 17, wherein the step of creating an annotation object includes ~~creating an annotation object~~ and storing the recorded audio input as part of the annotation object.

19. (Original) The method of claim 10, further comprising the step of comparing the audio input to a vocabulary to produce text.

20. (Currently Amended) The method of claim 19, wherein the step of creating an annotation object includes ~~creating an annotation object~~ and storing the text as part of the annotation object.

21. (Original) The method of claim 10, further comprising the steps of:
comparing the audio input to a vocabulary;
determining if the audio input has a matching entry in the vocabulary; and
storing the entry as part of the annotation object if the audio input has a matching entry in the vocabulary.

22. (Original) The method of claim 21, further comprising the steps of:
determining if the audio input has a close match in the vocabulary;
displaying the close matches;
receiving input selecting a close match; and
storing the selected close match as part of the annotation object if the audio input has
a close match in the vocabulary.

23. (Original) The method of claim 22, further comprising the step of displaying a message that the image has not been annotated if there is neither a matching entry in the vocabulary nor a close match in the vocabulary.

24. (Original) The method of claim 22, further comprising the following steps if there is neither a matching entry in the vocabulary nor a close match in the vocabulary:
receiving text input corresponding to the audio input;
updating the vocabulary with a new entry including the audio input and the text input;
and
wherein the received text is stored as part of the annotation object.

25. (Original) The method of claim 10, further comprising the steps of:
receiving text input corresponding to the audio input;
updating the vocabulary with a new entry including the audio input and the text input.

26. (Currently Amended) A computer implemented method for direct annotation of objects, the method comprising the steps of:

displaying an image;
receiving audio input;
detecting selection of [[an]] a location within the image;
comparing the audio input to a vocabulary to produce text; and
creating an annotation object, independent of the selected image, between the selected image location and the text, the annotation object including at least a text annotation field, an image reference field, and an annotation location field, the creating step occurring automatically in response to the receiving or detecting steps and including automatically terminating a recording of the audio input based on a predetermined audio level.

27. (Original) The method of claim 26, further comprising the step of recording the audio input received.

28. (Currently Amended) The method of claim 27, wherein the step of creating an annotation object includes creating an annotation object including a reference to the selected image location, the recorded audio input and the text, and storing the annotation object in an object storage.

29. (Currently Amended) The method of claim 26, wherein the step of creating an annotation object includes ~~creating an annotation object~~ and storing the text as part of the annotation object.

30. (Original) The method of claim 26, further comprising the steps of:

determining if the audio input has a matching entry in the vocabulary; and
storing the entry as part of the annotation object if the audio input has a matching
entry in the vocabulary.

31. (Currently Amended) The method of claim [[29]] 30, further comprising the steps
of:

determining if the audio input has a close match in the vocabulary;
displaying the close matches;
receiving input selecting a close match; and
storing the selected close match as part of the annotation object if the audio input has
a close match in the vocabulary.

32. (Currently Amended) The method of claim [[30]] 31, further comprising the step
of displaying a message that the image has not been annotated if there is neither a matching
entry in the vocabulary nor a close match in the vocabulary.

33. (Currently Amended) The method of claim [[30]] 31, further comprising the
following steps if there is neither a matching entry in the vocabulary nor a close match in the
vocabulary:

receiving text input corresponding to the audio input;
updating the vocabulary with a new entry including the audio input and the text input;
and
wherein the received text is stored as part of the annotation object.

34. (Previously Presented) The method of claim 26, further comprising the steps of:
receiving text input corresponding to the audio input;
updating the vocabulary with a new entry including the audio input and the text input.

35. (Currently Amended) A computer implemented method for displaying objects with annotations, the method comprising the steps of:
retrieving an image;
displaying the image with a visual notation that an annotation exists;
receiving user selection of ~~an image~~ the visual notation;
generating the annotation automatically, in response to user input of a location within
the image and an audio input, including automatically terminating a recording
of the audio input based on a predetermined audio level;
outputting ~~a notation~~ the annotation associated with the selected ~~image~~ visual
notation;
determining whether the annotation includes text;
retrieving a text annotation for the selected ~~image~~ visual notation; and
displaying the retrieved text with the image.

36. (Currently Amended) The method of claim 35, wherein the annotation is text and the step of outputting is displaying the text proximate [[an]] the image that it annotates.

37. (Previously Presented) The method of claim 35, wherein the annotation is an audio signal and the step of outputting is playing the audio signal.

38. (Canceled)

39. (Currently Amended) The method of claim 35, further comprising the steps of:
determining whether the annotation includes an audio signal;
retrieving [[a]] an audio signal for the selected image visual annotation; and
wherein the step of outputting is playing the audio signal.

40. (Currently Amended) A computer implemented method for retrieving images, the method comprising the steps of:

receiving audio input;
determining annotation objects that reference a close match to the audio input, each annotation object generated automatically in response to user input of a location within an image and an audio signal, where a recording of the audio signal is terminated automatically based on a predetermined audio level;
retrieving the images that are referenced by the determined annotation objects; and
displaying the retrieved images, the annotation object including at least an audio input field, an image reference field, and an annotation location field.

41. (Currently Amended) The method of claim 40, wherein the step of determining annotation objects further comprising comprises the steps of:

comparing the audio input to an audio signal reference [[by an]] of the annotation object; and

determining a close match between the audio input [[to]] and the audio signal reference [[by an]] of the annotation object if a probability metric is greater than a threshold of 80%.

42. (Currently Amended) The method of claim 40, wherein the step of determining annotation objects further ~~comprising~~ comprises the steps of:

determining the annotation objects for a plurality of images;

for each annotation object, comparing the audio input to an audio signal reference [[by an]] of the annotation object; and

determining a close match between the audio input [[to]] and the audio signal reference [[by an]] of the annotation object if a probability metric is greater than an a threshold of 80%.